

VOICING AFFECTS PERCEIVED MANNER OF ARTICULATION

*Fernández, S., Feijóo, S., Almeida, P.**

Departamento de Física Aplicada, Universidad de Santiago de Compostela
15706 Santiago de Compostela, SPAIN.
E-mail: fasergio@usc.es

*Instituto de Estudos Linguísticos, Universidade de Campinas
Campinas (SP), BRAZIL

ABSTRACT

The perception of voiced fricatives by native speakers of a language which lacks those phonemes is studied in this paper. Brazilian portuguese and Galician languages were chosen because they are historically related. A forced choice test reveals that listeners correctly perceive the place of articulation of the voiced fricatives. In order to examine whether the perception of fricative manner can be overridden by the voicing characteristics an open test was carried out. Listeners perceive voiced fricatives as a voiced phoneme with different manner of articulation and similar place of articulation or as its voiceless counterpart, depending on whether vocal-fold vibration extends over the whole obstruent interval or not. Results are discussed in terms of both historical phonetic changes and second language acquisition.

1. INTRODUCTION

Galician and Portuguese are two romance languages which evolved from the latin spoken in the Iberian Peninsula. During some time both languages belonged to a common branch called Galician-Portuguese language, which later split when Portugal became a separate Kingdom. One of the main differences between their phonetic systems, as spoken in Brazil and Galicia (Spain), is that Galician language lacks voiced fricatives.

Brazilian portuguese fricatives have the same three places of articulation for both the unvoiced (/f/, /s/, /ʃ/) and voiced (/v/, /z/, /ʒ/) fricatives. The characteristics of Portuguese voiced fricatives are similar to the corresponding English voiced fricatives, which have been studied in a number of works (see for instance Baum & Blumstein, 1987; Stevens *et al.*, 1992). Galician lacks voiced fricatives and has an additional place of articulation (/θ/). Voiced Galician fricatives may have been lost during its independent evolution, although that question still remains unclear (Fradejas, 97; Ariza, 94). Nevertheless, voiced phonemes with a manner of articulation different from the portuguese fricatives and with similar place of articulation, may be found in Galician language : /b/ (bilabial stop) for /v/ (labiodental fricative), and /j/ (voiced alveolar affricate) for /ʒ/ (alveolar fricative). The portuguese phoneme /z/ lacks a voiced cognate in Galician language.

The purpose of this paper is to study the perception of the Brazilian portuguese voiced fricatives by Galician speaking subjects, in order to determine how the listeners of a language which lacks voiced fricatives perceive those voiced fricatives. Up till now,

most cross-language studies have been devoted to the perception of vowels (for studies involving Spanish, see for instance Cutler *et al.*, 1996; Flege & Munro, 1994; Flege, 1991).

Two hypothesis can be considered: 1) Listeners would identify the voiced fricatives as their unvoiced cognates; 2) Listeners may perceive some of the voiced fricatives as some other voiced consonants.

In order to test the above hypothesis, closed and open perceptual tests were carried out. The closed test will give us an indication of whether the galician listeners are able to correctly perceive the place of articulation of the voiced fricatives, while the open test will show whether the perception of fricative manner can be overridden by the voicing characteristics of the portuguese voiced fricatives.

2. MATERIALS AND METHOD

The tokens, which were pronounced into a carrier phrase with meaning in both languages (“*digo* CV-syllable *baixinho*”), correspond to CV syllables formed by the combination of the three fricatives /z, v, ʒ/ with each of the eight brasilian vowels /a, e, ε, i, ɔ, o, u/. Tokens were pronounced by one man and one woman. Stevens *et al.* (1992) found that American speakers produced the voiced fricatives with two forms of voicing: 1) Extended form: vocal-fold vibration extends over the whole obstruent interval, and 2) Non extended form: vocal-fold vibration does not overlap with frication noise during the whole obstruent interval. Inspection of the tokens revealed that the male speaker produced the fricatives with the extended form of voicing, while most of the tokens produced by the female speaker corresponded to the non-extended form. Tokens of this second form must be considered with caution, particularly some tokens with /v/ were considered by the authors to overlap the characteristics of the two groups. Both forms of voicing can be seen in figures 1 and 2. Thus, the total number of tokens is: 42 tokens = 2 forms of voicing × 3 fricatives × 7 vowels.

Tokens were recorded in a soundproof room located in the Instituto de Estudos Linguísticos (IEL) of the Universidade de Campinas (Brazil), using a DAT recorder, and later digitized with a frequency of 16 kHz using the CSL system (Computer Speech Laboratory). The perceptual tests were carried out with a PC-486 computer in the Facultad de Física of the Universidad de Santiago.

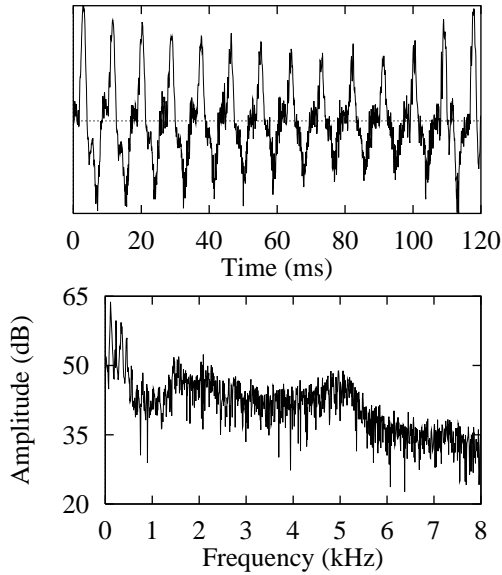


Figure 1: Fricative noises for the extended form of voicing. Top: temporal signal, bottom: spectrum.

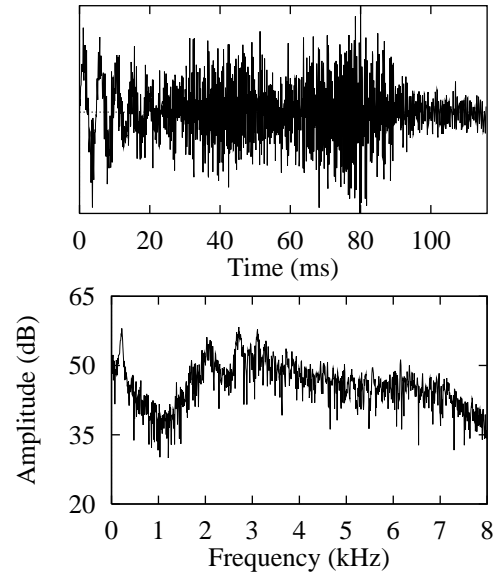


Figure 2: Fricative noises for the non extended form of voicing. Top: temporal signal, bottom: spectrum.

Visual and auditory inspection along with spectrograms of the tokens were used to isolate fricative noises and syllables. For the isolation of the fricative noises the spectrogram was plotted and the segment of the consonant which clearly contained high frequency noise was selected as the fricative noise segment.

3. PERCEPTUAL EXPERIMENTS

Twenty university students with normal hearing, who were native speakers of galician, served as subjects for the perceptual experiments. They volunteered to participate in the experiments for course credit. Three conditions are considered in the perceptual experiments. Condition one corresponds to the whole phrase, condition two corresponds to the syllable and condition three corresponds to the fricative. Two test were performed with the 126 tokens ($3 \text{ fricatives} \times 7 \text{ vowels} \times 2 \text{ forms of voicing} \times 3 \text{ conditions}$): I) Forced choice test: in order to study confusions of place of articulation, a three alternative ($/s, f, \text{f}/$) forced choice was performed; II) Open test, with $/s, f, \text{f}, b, j/$ and *other* as options, in order to assess the importance of voicing versus manner of articulation cues. Because of historical phonetic changes and the existence of allophones, there is no real agreement about which fricative or affricate phoneme corresponds to the Galician language sound that is included here as option $/j/$, since some times that sound is produced as the lateral $/\lambda/$.

Perceptual experiments were carried out in a normal office at the Faculty of Physics using a computer program developed at our laboratory, which controls the whole process. The forced choice test was carried out prior to the open test. Each test consisted of the following steps: first, the meaning of the test was explained to the subjects; second, some tokens of each condition were presented to the subjects in order to familiarize them with the sounds they were going to hear, not to train them for that task; and third, the test was carried out. Subjects were presented the

stimuli through SONY MDR-570 headphones. One repetition of each token was allowed, after which it was mandatory to select one of the possible options.

3.1. Results

The identification percents for the forced choice test were: 94.5% in the phrase condition, 95.6% in the syllable condition and 88.8% in the fricative condition. Confusion matrices for the forced choice test can be seen in Table 1, identification of the place of articulation for the fricative $/v/$ being almost perfect: 100.0% for the phrase condition, 99.3% for the syllable condition and 96.4% for the fricative condition. Identification of the place of articulation for $/z, \text{z}/$ is satisfactory either for the phrase (88.9% and 94.6%, respectively), syllable (92.9% vs. 94.6%) or fricative conditions (82.5% vs. 87.5%).

An analysis of variance on the correct responses revealed a significant effect for condition ($F(2, 123) = 8.2, p < 0.0005$), identification of place in the fricative condition being significantly lower than in the phrase and syllable conditions. Also, a significant effect for fricative showed up ($F(2, 123) = 20.0, p < 0.00005$), identification being significantly different for each fricative: place of articulation of $/z/$ was the worse identified, whereas place of articulation of $/v/$ was the best identified. No significant effect for form of voicing showed up.

Confusion matrices for the open test are shown in Tables 2, 3 and 4 for the phrase, syllable and fricative condition respectively. The percent of voiceless responses in the phrase condition was 35.2% for the extended form of voicing and 62.6% for the non-extended form of voicing. In the syllable condition they were 45.0% and 73.6% respectively, and in the fricative condition, 45.7% and 66.9%. This indicates similar results for the three conditions and that the number of voiceless responses is larger for the non-extended form than for the extended form. Note that

	/f/	/s/	/ʃ/
/v/	100.0	0.0	0.0
/z/	1.4	88.9	9.6
/ʒ/	0.4	5.0	94.6
/v/	99.3	0.7	0.0
/z/	0.7	92.9	6.4
/ʒ/	0.7	4.6	94.6
/v/	96.4	2.5	1.1
/z/	7.5	82.5	10.0
/ʒ/	1.1	11.4	87.5

Table 1: Confusion matrices for the forced choice test. Top: phrase condition, middle: syllable condition, bottom: fricative condition

	/f/	/b/	/s/	/ʃ/	/j/	other
/v/	27.1	68.6	0.0	0.0	0.0	4.3
/z/	0.7	0.7	62.2	9.3	5.7	21.4
/ʒ/	0.0	0.0	3.6	16.4	77.9	2.1
/v/	36.4	61.4	0.0	0.0	0.0	2.2
/z/	0.0	0.0	85.7	8.6	2.9	2.8
/ʒ/	0.0	0.0	3.6	65.7	30.0	0.7

Table 2: Confusion matrices for the open test in the phrase condition. Top: *extended* form of voicing, bottom: *non-extended* form of voicing.

	/f/	/b/	/s/	/ʃ/	/j/	other
/v/	37.1	54.3	0.0	0.7	0.0	7.9
/z/	2.9	0.7	75.7	6.4	1.4	12.9
/ʒ/	0.0	0.0	3.6	22.1	69.3	5.0
/v/	47.1	50.7	0.0	0.0	0.0	2.1
/z/	0.0	0.0	87.2	5.0	1.4	6.4
/ʒ/	0.7	0.0	2.9	86.4	8.6	1.4

Table 3: Confusion matrices for the open test in the syllable condition. Top: *extended* form of voicing, bottom: *non-extended* form of voicing.

	/f/	/b/	/s/	/ʃ/	/j/	other
/v/	49.3	43.6	0.7	0.7	2.1	3.6
/z/	6.4	5.7	57.2	8.6	0.7	21.4
/ʒ/	0.0	1.4	5.7	30.7	53.6	8.6
/v/	52.9	27.9	1.4	0.7	0.0	17.1
/z/	0.7	1.4	77.2	6.4	0.7	13.6
/ʒ/	1.4	0.7	5.7	70.7	9.3	12.2

Table 4: Confusion matrices for the open test in the fricative condition. Top: *extended* form of voicing, bottom: *non-extended* form of voicing.

for the extended form /z/ which does not have a voiced cognate is identified as *other* in the 21.4% of the cases in the phrase and fricative conditions and in the 12.9% of the cases in the syllable condition.

For the open test, a significant main effect for form of voicing ($F(1, 124) = 56.8, p < 0.0005$) showed up, phonemes being more often identified as its voiceless cognate for the non-extended form of voicing than for the extended form of voicing. The effect of fricative ($F(2, 123) = 33.5, p < 0.0005$) was also significant: a Scheffé test revealed that /z/ was significantly more identified as its voiceless cognate (/s/) than /v/ or /ʒ/. There were no significant differences among the three conditions.

2-way fricative \times form of voicing interaction was also significant ($F(2, 123) = 14.8, p < 0.0005$). In order to study this interaction, two separate ONEWAY analysis of variance on correct responses were carried out. First, the effect of the form of voicing factor was only significant for /z/ ($F(1, 40) = 13.6, p < 0.0007$) and /ʒ/ ($F(1, 40) = 76.8, p < 0.00005$). Second, the effect of the fricative factor was significant both for the extended ($F(2, 60) = 25.7, p < 0.00005$) and the non-extended ($F(2, 60) = 30.4, p < 0.00005$) forms of voicing. For the extended form of voicing the /z/ phoneme was significantly more identified as its voiceless cognate (/s/) than /v/ and /ʒ/. This effect does not showed up in the fricative condition where the identification of the three fricatives was not significantly different. For the non-extended form of voicing the /z/ and /ʒ/ phonemes were significantly more identified as their voiceless cognates (/s/ and /ʃ/, respectively) than /v/. This effect was not significant for the fricative condition where only the identification of /v/ and /z/ was significantly different. The /v/ fricative was quite well identified as /b/ in the non-extended form of voicing which might be a consequence of the difficulty of pronouncing this particular fricative in the non-extended form as it was already noted.

The analysis of the perceptual experiments show that subjects identify almost perfectly the place of articulation of voiced fricatives which they are not accustomed to perceive. Thus, place of articulation confusion does not explain the results of the open test. In fact, they may perceive a certain stimuli as being a voiced phoneme with different manner of articulation, as a voiceless phoneme with the same manner of articulation, depending on the form of voicing with which the token was pronounced. In the

case of /v/, listeners may even perceive a slightly different place of articulation (/b/). For the extended form, voicing is a strong cue and /z/, which lacks a voiced cognate, is identified often as *other* phoneme.

4. DISCUSSION

The results of the forced choice tests indicate that Galician listeners are able to perceive the correct place of articulation in Portuguese voiced fricatives. /v/ is the voiced fricative with the place of articulation most easily recognized by the listeners, while there are some confusions between the places of articulation of /z/ and /ʒ/. Thus, /v/ benefits from the absence of a voiced cognate of the fricative /θ/ in Portuguese, since Galician listeners tend to confuse /f/ and /θ/ quite frequently (Feijóo *et al.*, 1998).

The open tests show a somewhat different picture. The first notable effect is a significant difference between the two forms of voicing. Stevens *et al.* (1992) found in their work on English fricatives voicing that there was a great deal of variation in the production of voicing in fricatives, listeners being sensitive in different degrees to those differences. In our study, this effect is stronger in the fricative and syllable conditions, since in those conditions voicing prior to fricative onset was removed and, as a consequence, more voiceless responses are given to the non-extended form stimuli, since their fricative noises have very little (if any) voicing overlapped with the friction noise. Then, while in the extended form /ʒ/ tends to be identified as /j/ (both of them are voiced phonemes), in the non-extended form, it tends to be identified as its voiceless cognate /ʃ/. For /z/, the picture is a little different since the phonetic inventory of the listeners lacks a voiced phoneme with the same place of articulation. As a result, in the extended form, where voicing overlaps with the fricative, /z/ is generally recognized as /s/, but there are a considerable number of “other” responses, indicating that listeners perceived voicing in this case as a extraneous sound. It is also interesting to observe that /z/ in the extended form produced quite a number of confusions with other places of articulation in a higher degree than /v/ or /ʒ/, which in some way confirms the above conclusion. In the non-extended form /z/ was usually identified as /s/. The case with /v/ is somewhat different from those above. First, there was very little difference between the results of both forms of voicing. Second, the difference between the voiced and voiceless responses was in general less marked than for the case of /ʒ/, except in the phrase condition. Listeners showed some preferences for the /b/ responses, though.

These results can be analysed taking into consideration the probable historic changes in sound and phonetic inventory that took place from the Middle to the Modern ages in the Iberian Peninsula. According to Fradejas (1987), /s/ and /z/ coexisted during sometime and later merged into a single /s/ phoneme, probably due to the lesser effort needed to produce the voiceless fricative. Results of the open and closed tests confirm that Galician listeners associate /z/ only to /s/, since there is no other phoneme with close acoustic characteristics. For /ʒ/ the situation is more complicated, since there is no complete agreement as to the changes and the directions in which those changes took place. /ʒ/ seems to have lost its voicing characteristics and transformed into /ʃ/. On the other hand, /ʒ/ is thought to have come from a /ʎ/ (lateral palatal), which is produced in the Galician region as a voiced affricate /dʒ/, by a process of acquisition of fricative characteristics. Our re-

sults confirm the perceptual relation between the three phonemes, which differ among them in voicing and manner of articulation. The case of /v/ is more obscure, since there is no agreement as to whether that phoneme really existed in the Iberian Peninsula (Fradejas, 1997; Ariza, 1994). According to some accounts, /f/ was transformed into the bilabial fricative /β/, which later turned into /b/. But, since /f/ is labiodental, there is still the possibility of /f/ having turned into /v/, which later would turn into /b/. Our results suggest a strong relationship between the perceptual characteristics of /v/ and /b/.

Actual theories about second language acquisition state that linguistic experience plays a substantial role in the perception of consonants (Iverson & Kuhl, 1996). Listeners from L1 assimilate new phonemes of L2 to their native L1 space. Our results show that this is indeed the case: Although the listeners were instructed to choose the “other” option whenever they felt that the sound they heard did not correspond to any of the possible choices, most of them chose phonemes of their L1 inventory. Their choices resulted sometimes in a change of manner or place, sometimes in a change of voicing characteristics, depending upon the phoneme and the form of voicing (extended, non-extended). So, the effects of L1 on the perception of L2 depend not only on the L1 inventory, but also on the production of L2 tokens, at least in the case of phonemes like voiced fricatives, which can be produced in different ways depending on the speaker style.

5. REFERENCES

1. Ariza, M., “Sobre fonética histórica del español,” *Ed. Arco*, Madrid, 1994.
2. Baum, S.R., Blumstein, S.E., “Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English,” *J. Acoust. Soc. Am.*, vol. 82, pp. 1073–1077, 1987.
3. Cutler, A., van Ooijen, B., Norris, D., Sánchez-Casas, R., “Speeded detection of vowels: A cross-linguistic study,” *Perception and Psychophysics*, vol. 58, pp. 807–822, 1996.
4. Feijóo, S., Fernández, S., Balsa, R., “Context effects in the auditory identification of the Spanish fricatives /f/ and /θ/: Hyper and Hypospeech,” *Proceedings of the ICA/ASA '98*, pp. 2025–2026, 1998.
5. Flege, J.E., Munro, M.J., Fox, R.A., “Auditory and categorical effects on cross-language vowel perception,” *J. Acoust. Soc. Am.*, vol. 95, pp. 3623–3641, 1994.
6. Flege, J.E., “The interlingual identification of Spanish and English vowels: Orthographic evidence,” *Quarterly Journal of Experimental Psychology*, vol. 43A, pp. 701–731, 1991.
7. Fradejas, J.M., “Fonología histórica del español,” *Ed. Visor libros*, Madrid, 1997.
8. Iverson, P., Kuhl, P.K., “Influences of phonetic identification and category goodness on American listeners perception of /t/ and /l/,” *J. Acoust. Soc. Am.*, vol. 99, pp. 1130–1140, 1996.
9. Stevens, K.N., Blumstein, S.E., Glicksman, L., Burton, M., Kurowski, K., “Acoustical and perceptual characteristics of voicing in fricatives and fricative clusters,” *J. Acoust. Soc. Am.*, vol. 91, pp. 2979–3000, 1992.